CLAIMS

What is claimed is:

1. A method of media editing, comprising:

receiving audio data and a plurality of associated audio descriptors, which describe characteristic of said audio data;

receiving visual data and a plurality of associated visual descriptors, which describe characteristic of said visual data;

determining a plurality of corresponding weights for said visual data;

correlating said audio data and said visual data based on said corresponding weights, said associated audio descriptors, and said associated visual descriptors; and

adjusting said audio data and said visual data to construct a media output.

2. The method of media editing according to claim 1, further comprising rendering said media output with style information.

3. The method of media editing according to claim 1, wherein the step of receiving audio data and said associated audio descriptors comprises:

receiving an audio signal; and

analyzing and segmenting said audio signal for generating said audio data and said associated audio descriptors, wherein said audio data consists of a plurality of audio segments.

4. The method of media editing according to claim 1, wherein the step of receiving visual data and said associated visual descriptors comprises receiving a plurality of visual segments and said associated visual descriptors.

5. The method of media editing according to claim 4, wherein the step of determining a plurality of corresponding weights comprises calculating any said corresponding weight for respective said visual segment.

6. The method of media editing according to claim 5, wherein the step of correlating comprises:

extracting an audio duration, from said associated audio descriptors, for respective said audio segment ;

extracting a visual duration, from said associated visual descriptors, for respective said visual segment;

evaluating a plurality of correlating scores for respective sequences of said visual

11

segments, based on said corresponding weights, said corresponding audio durations and said corresponding visual durations; and

finding a sequence of visual segments with a correlating score that is the maximal within said plurality of correlating scores.

7. The method of media editing according to claim 4, wherein the step of receiving audio data and said associated audio descriptors comprises:

receiving an audio signal; and

generating a plurality of audio indices by choosing said audio signal with audio change therein.

8. The method of media editing according to claim 7, wherein the step of correlating comprises:

finding a duration on each said visual segment;

determining a searching window based on said duration;

finding, within said searching window, a first index on said audio indices, wherein said first index is more than other indices on said audio indices within said searching window; and

adjusting each said visual segment, based on a time corresponding to said first index.

9. The production method of media output, comprising:

receiving audio segments and a plurality of associated audio descriptors, which describe characteristic of said audio segments;

receiving visual segments and a plurality of associated visual descriptors, which describe characteristic of said visual segments;

determining a plurality of corresponding weights for each said visual segment;

extracting a visual duration, from said associated visual descriptors, for each said visual segment;

extracting an audio duration, from said associated audio descriptors, for each said audio segment;

evaluating a plurality of correlating scores for respective sequences of said visual segments, based on said corresponding weights, said corresponding audio durations and said corresponding visual durations;

finding a sequence of visual segments with a correlating score that is the maximal within said plurality of correlating scores; and

adjusting said audio segments and said visual segments to generate a media output.

10. The production method of media output according to claim 9, further comprising rendering said media output with style information.

11. The production method of media output according to claim 9, wherein the step of receiving audio segments and associated audio descriptors comprises:

receiving an audio signal; and

analyzing and segmenting said audio signal for generating said audio segments and said associated audio descriptors.

12. The production method of media output according to claim 9, wherein the step of receiving visual segments and associated visual descriptors comprises:

receiving an video signal; and

analyzing and segmenting said video signal for generating said video segments and said associated visual descriptors.

13. The production method of media output according to claim 9, wherein said visual segments and said associated visual descriptors are in format of MPEG-7.

14. The production method of media output according to claim 9, wherein said audio segments and said associated audio descriptors are in format of MPEG-7.

15. The production method of media output, comprising:

receiving audio data and a plurality of associated audio descriptors, which describe characteristic of said audio data;

receiving visual data and a plurality of associated visual descriptors, which describe characteristic of said visual data;

finding, within a searching window, a value corresponding to said associated audio descriptors on said audio data, wherein said value is more than other value corresponding to associated audio descriptors within said searching window; and

adjusting said visual data, based on a time corresponding to said value, to generate a media output, wherein said media output is based on audio data and said adjusted visual data.

16 The production method of media output according to claim 15, further comprising rendering said media output with style information.

17. The production method of media output according to claim 15, wherein said visual data and said associated visual descriptors are in format of MPEG-7.

18. The production method of media output according to claim 15, wherein said audio data and said associated audio descriptors are in format of MPEG-7.

19. The production method of media output according to claim 15, wherein the step of receiving said audio data and said associated audio descriptors comprises:
    receiving an audio signal; and
    generating a plurality of audio indices by choosing said audio signal with audio change therein.

20. A storage device, storing a plurality of programs readable by a media process device, wherein the media process device according to said programs executes the steps comprising:
    receiving audio data and a plurality of associated audio descriptors, which describe characteristic of said audio data;
    receiving visual data and a plurality of associated visual descriptors, which describe characteristic of said visual data;
    determining a plurality of corresponding weights for said visual data;
    correlating said audio data and said visual data based on said corresponding weights, said associated audio descriptors, and said associated visual descriptors; and
    adjusting said audio data and said visual data to construct a media output.

21. A storage device, storing a plurality of programs readable by a media process device, wherein the media process device according to said programs executes the steps comprising:
    receiving audio segments and a plurality of associated audio descriptors, which describe characteristic of said audio segments;
    receiving visual segments and a plurality of associated visual descriptors, which describe characteristic of said visual segments;
    determining a corresponding weight for each said visual segment;
    extracting a visual duration, from said associated visual descriptors, for each said visual segment;
    extracting an audio duration, from said associated audio descriptors, for each said audio segment;
    evaluating a plurality of correlating scores for respective sequences of said visual segments, based on said corresponding weights, said corresponding visual durations and said corresponding audio duration;
    finding a sequence of visual segments with a correlating score that is the maximal within said plurality of correlating scores; and

adjusting said audio segments and said visual segments to generate a media output.

22. A storage device, storing a plurality of programs readable by a media process device,
5   wherein the media process device according to said programs executes the steps comprising:

receiving audio data and a plurality of associated audio descriptors, which describe characteristic of said audio data;

receiving visual data and a plurality of associated visual descriptors, which
10   describes characteristic of said visual data;

finding, within a searching window, a value corresponding to said associated audio descriptors on said audio data, wherein said value is more than other value corresponding to said associated audio descriptors within said searching window; and

adjusting said visual data, based on a time corresponding to said value, to
15   generate a media output, wherein said media output is based on audio data and said adjusted visual data.